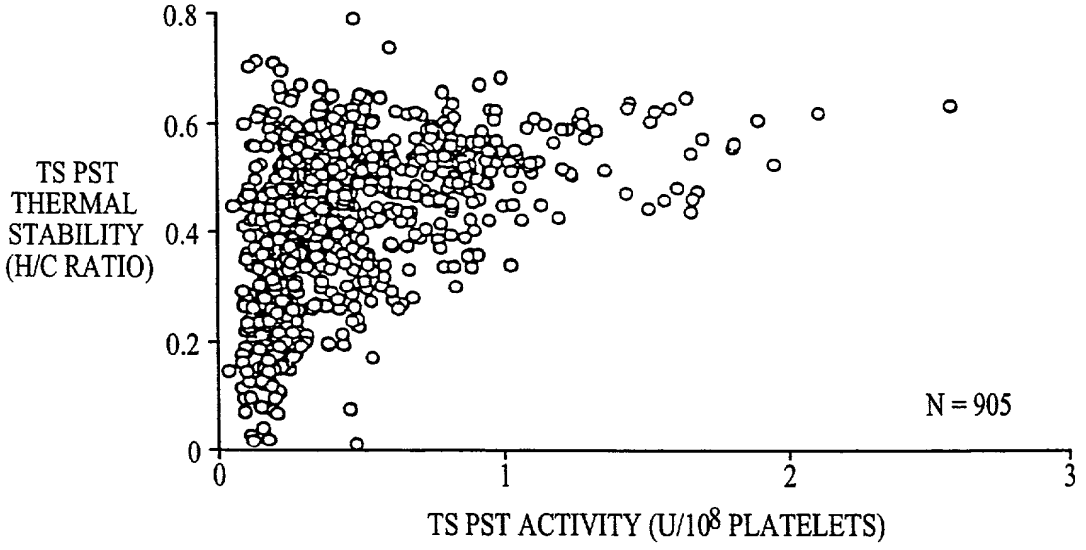
0.8
0.6
0.4
0.2
0
TS PST
THERMAL
STABILITY
(H/C RATIO)
N = 905
0
1
2
3
TS PST ACTIVITY (U/10^8 PLATELETS)

FIG. 1A

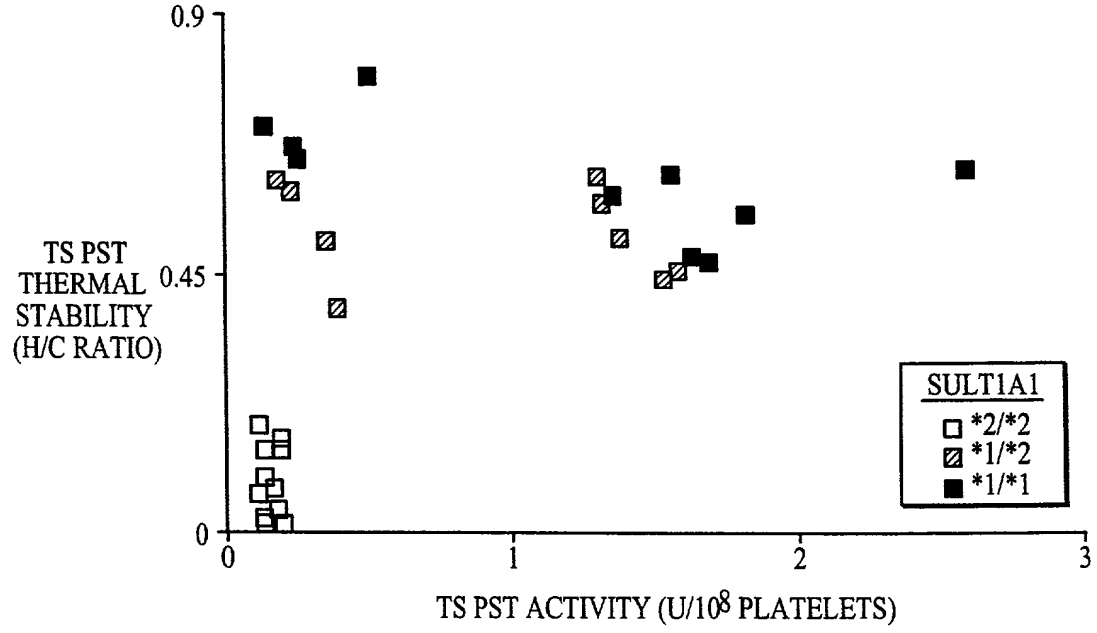
0.9
0.45
0
TS PST
THERMAL
STABILITY
(H/C RATIO)
SULT1A1
*2/*2
*1/*2
*1/*1
0
1
2
3
TS PST ACTIVITY (U/10^8 PLATELETS)

FIG. 1B

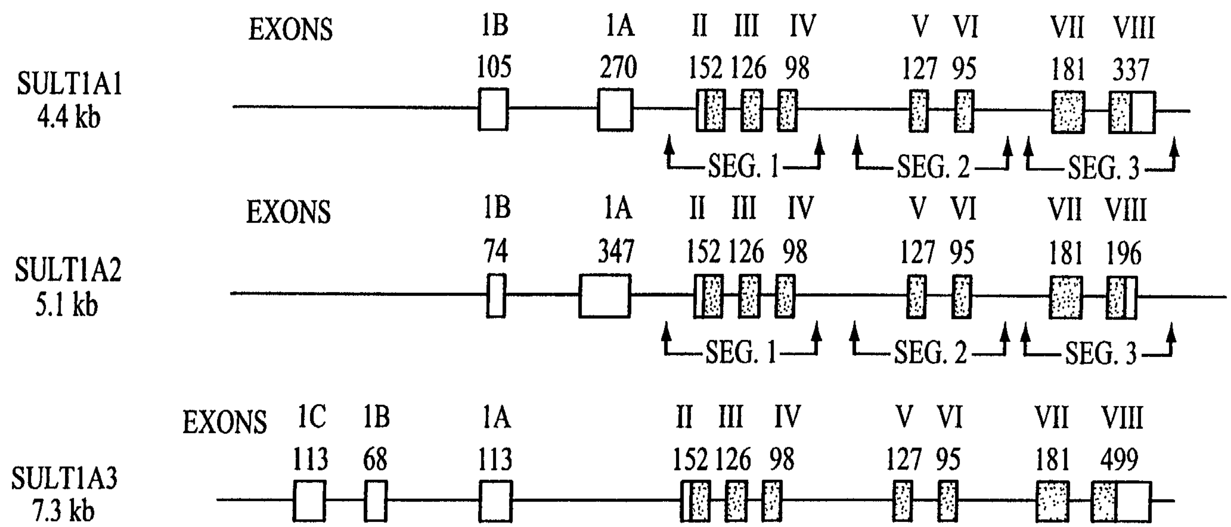

FIG. 2

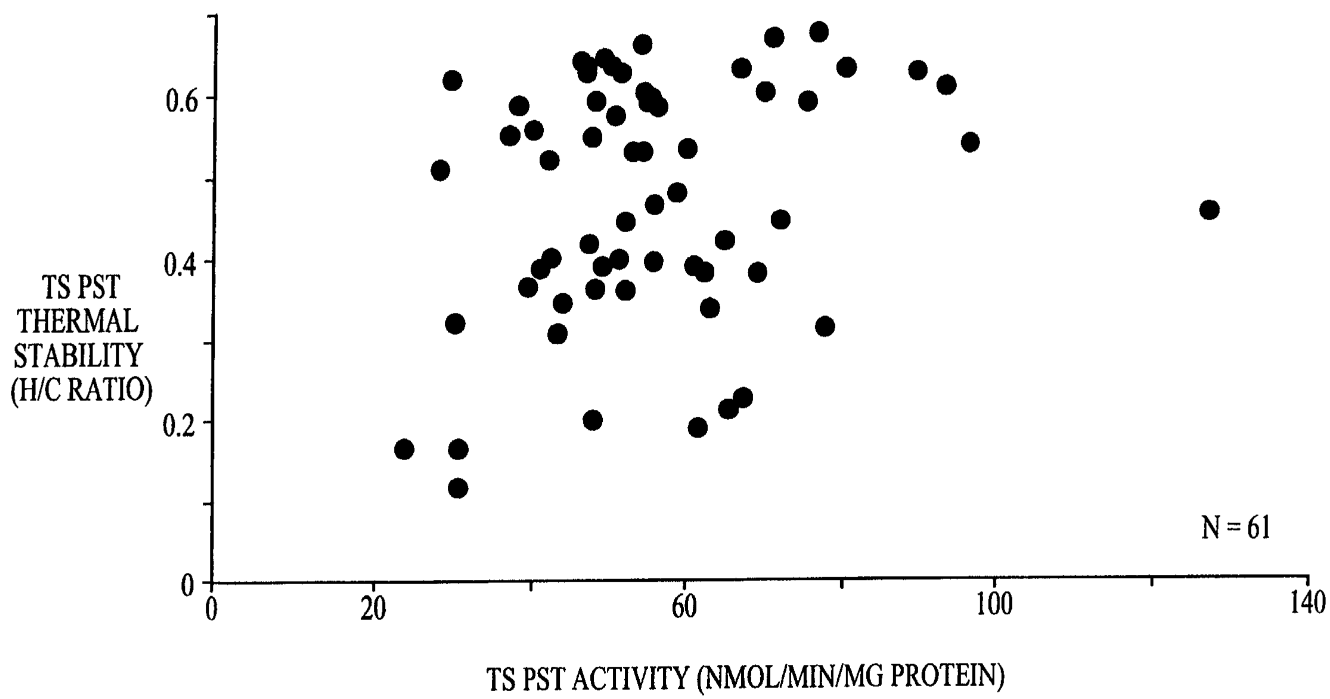
TS PST
THERMAL
STABILITY
(H/C RATIO)
0.6
0.4
0.2
0
0
20
60
100
140
N = 61
TS PST ACTIVITY (NMOL/MIN/MG PROTEIN)

FIG. 3

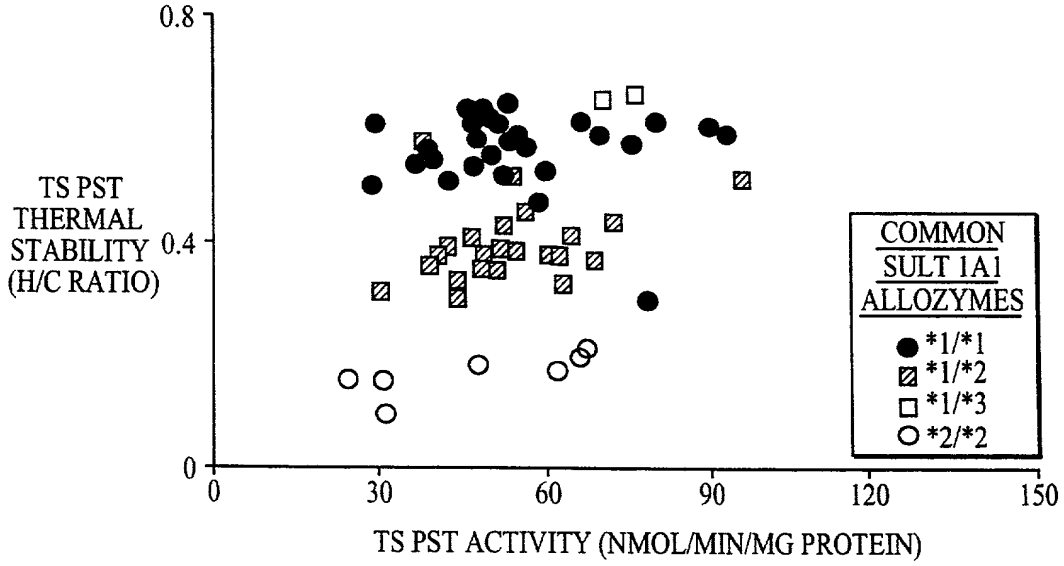
0.8
0.4
0
TS PST
THERMAL
STABILITY
(H/C RATIO)
0
30
60
90
120
150
TS PST ACTIVITY (NMOL/MIN/MG PROTEIN)
COMMON
SULT 1A1
ALLOZYMES
*1/*1
*1/*2
*1/*3
*2/*2

FIG. 4A

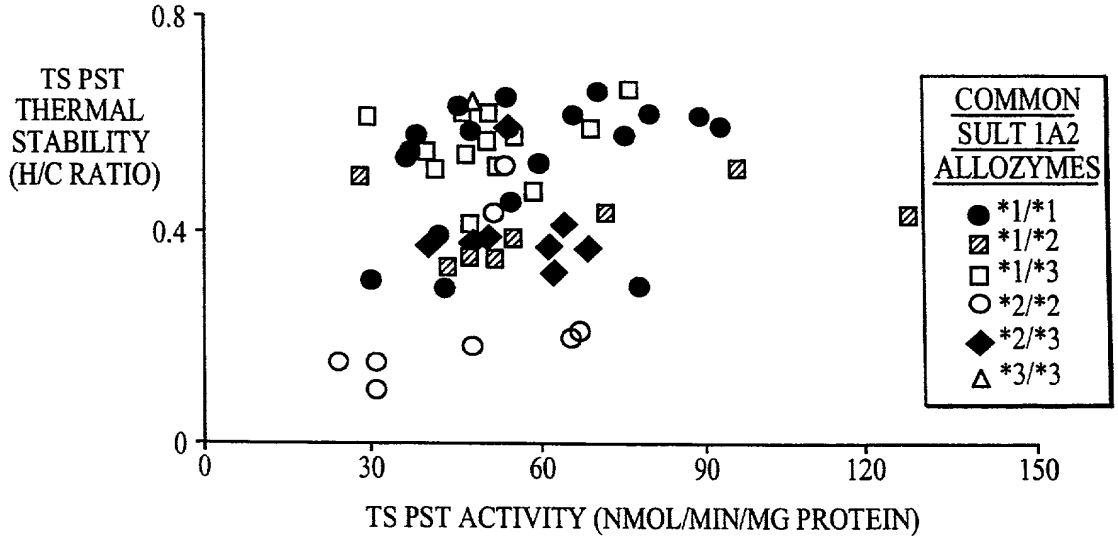
0.8
0.4
0
TS PST
THERMAL
STABILITY
(H/C RATIO)
0
30
60
90
120
150
TS PST ACTIVITY (NMOL/MIN/MG PROTEIN)
COMMON
SULT 1A2
ALLOZYMES
*1/*1
*1/*2
*1/*3
*2/*2
*2/*3
*3/*3

FIG. 4B

```
      ttgctgccag ctgcctctcc ctccttgtct cttacctgcc tgctgcctgg gacaggatga agcggggccc ttgtgttgcc ccaaccctgg ctgttggcta
101   agagcccacg tgatctgcct gtgagaggag ttccttccgg aagaaccagg gcagcttctg ccctagagg gccaatgccc tagctgagtg cagtcccccg
201   gccccagcct ggtccagctt tgggaagagg gtgcccagtt gtgcaatcca ggccggggca gccgtgtcct gatcttggta ttcagggctg agcctggagg
301   gggcttgtga tgcctgactc tgtctccctc tctggcccca tgccttggta gctgtgaggc gtcactgctt tgggtgacct gatctggctg tgatggatga
401   gcacgggga aatagtggaa gactcggaat tagaagacgt gagtgggctt tggccccagc ctccctaccc cactccctgt cctgggctgc ctgtgaccaa
501   cctcgtttct gcaggcacac tggatagccc tgctggagct cagtgtccct aatcccctcc agatactggt ggcctagggg aggtcatcaa aaaccggtgg
601   gacatcgacc tcagcccgtt tccacgcttt tttttgtttt tttttttttt ttgagaccga gtttcactct tgttgcccag gctggagtgc aatggcgtga
701   tcttggctca ccgcaacctc cgcctcctgg gttcaagcga ttctcctgcc tcagcctccc aagtagctgg gattccaggc gtgtgccacc aggcttgact
801   aattttctat ttttagtaga gacaaggttt ctccatgttg gtcaggctgg tctcaaactc ccgacttcag gtgatctgcc tgcctcggcc ccccaaagtg
901   ctgggattac aggagtgagc caccgtgcca ggccttctcc aggctcttgg caccttagcc agaaacaatt taaggacaag tgcaaaagtc atgaatgtag
1001  gcagatttcc tgcagagtaa agggactcac tcaagaagag gaacgtgggg gtcctcaaga gagtgtctca tgccctacaa ggtgtggggc tgacctttat
1101  gggcttcttc aactaaagag gggtatattc atgaagagtc caggaaaagg taaagatttc tcaagaccgt ggtgccacaa tttacaccca aatacaggtg
1201  ttcctggagc cgtcttggca ctggtgggtg tacggtttca tatgttactg atcatacaat gagatcctag gtgaaaccta catcaaatac agcgccatgt
1301  tgtgtctggt tggtcgtggc cagcttggtc ctcatcctat ttttcaggga cttattggcc cttagcgcat gcagctattt caagtttcct tcttctcctc
1401  atgtgaaact gctgcctggg attttgtatt cacttgctac cactctatta atctcacatt ctcgcctctt ttctgtgtca ccccgtgtgg gtccgacagg
1501  ttgttactag agtgcaatac aaagtcttag tcaagggaac ctcctgaggg ttgctgaggg caggggtgga gctagtagcc tgaggacctg ccagtcacgg
1601  ggattcctca tgggcacaga ggagggagga ggggtccatg gccctagcat atgagaagcc tctcctctgc ctggaattcc catgcctcag cttccccccac
1701  actcccacct gtccgcttgc ctctgaactc acgcatttct tggaagtctt gggagattca cctttactca gatggttgtt tacctgtctc gtgcacagct
1801  tgaccttgga ctttaaagtg aggataaaga acgaggagga tgggggatg cccccttcc acgggccctg tggcttccaa acctcggcct cctctggtct
1901  cttgtctgtg gagcctcctt caaacccagg gaaataaaac cacctgccac gggttgtggt tcttctagga tcttctatca atgttctctg aggtccccag
2001  gagccatgaa gctggggctg actcccaggg caatgggact gcagtgtcct tgttctttct tgttctatgc atccatgctc tgctccaccc ctgccccttc
2101  actctgccca cacacatccc tctagactgg ccttgtggtc agagcctgga gtgcatgggc tgctgggggc ctgtgggctg cactgggcca gaacccctgg
2201  caccttcaag actggcctgg agccagcagg taggtgacct ttccagggcc tgcctatccc agctttctcc tccaatccct cccctctctt gcctgggtca
2301  attagagaga gcttgtctgt tggctgcctg gcgggtgga gttcaggggc aggtcaggag cccagtgaca gctcggaaaa aaaaaaaaaa aaaaaaaaaa
2401  cagaaaaaaa aacctacaaa aacaaaccca ccattgggcc tttccccttt cattcttctg ttttctacac agcaaactca gtcgtggctt tggagatcac
2501  tttaagcttg tctccagctg gcacactaag gagggtaatg gagaagctcc cccaccccca accccacccc ttccttccgg aagcaaatct aagtccagcc
```

FIG. 5A

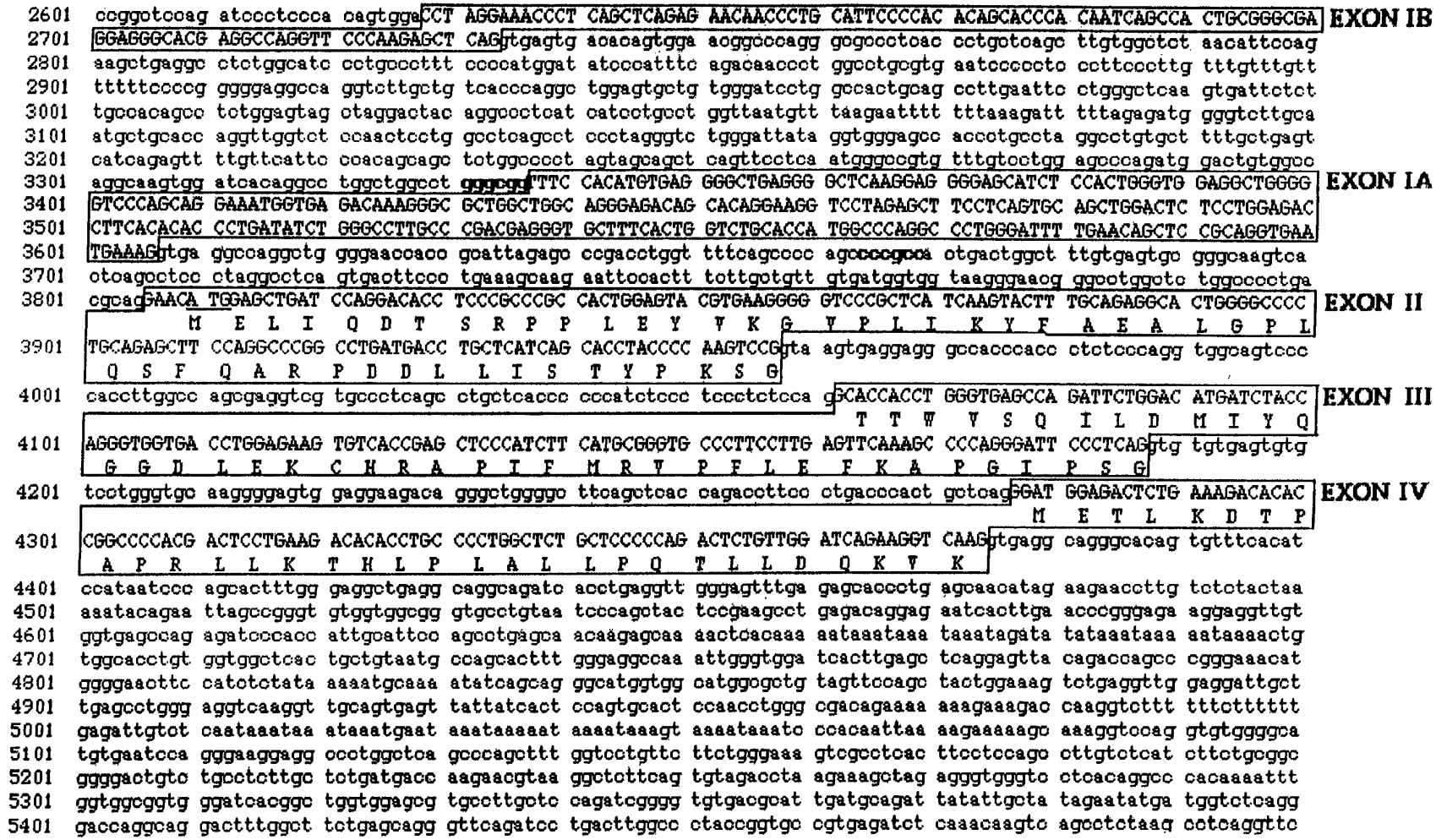

FIG. 5B

```
5501  ctcctttgcc aatccaagag atgagctggc ctggggcagg ctgtgtggtg atggtgctgg ggttgagtct tctgcccctg cagGTGGTCT ATGTTGCCCG   EXON V
                                                                                                 V  V  Y   V  A  R
5601  CAACGCAAAG GATGTGGCAG TTTCCTACTA CCACTTCTAC CACATGGCCA AGGTGCACCC TGAGCCTGGG ACCTGGGACA GCTTCCTGGA GAAGTTCATG
       N  A  K   D  V  A  V   S  Y  Y   H  F  Y   H  M  A  K   V  H  P   E  P  G   T  W  D  S   F  L  E   K  F  M
5701  GTCGGAGAAG gtgggtttga tgggaggaag gaaagtgtgg agccgagggg tggtggctac aacgcacagc aaccctgtgt tggcaccoct tgcctgcttc
      V  G  E  V
5801  tccagTGTCC TACGGATCCT GGTACCAGCA CGTGCAGGAG TGGTGGGAGC TGAGCCGCAC CCACCCTGTT CTCTACCTCT TCTATGAAGA CATGAAGGAG   EXON VI
         S   Y  G  S  W   Y  Q  H   V  Q  E   W  W  E  L   S  R  T   H  P  V   L  Y  L  F   Y  E  D   M  K  E
5901  gtgagaccac ctgtgaagct tccctccatg tgacacctgg gggccggcac ctcacaggga cccaccaggg tcacccagcc ccctcccttg gcagcccccca
6001  cagcaggccc ggattcccca tcctgccttc ttggcccagg cctccccgct acaggcccca cctggcagcg ggccccacac ggctctcatc accccacatct
6101  gagtcagctg catggggggc cacggatcag aaacttagtc ctattgctac tccctgccaa agggtgtgcc acccagggcc acagtcatgg aagaagacca
6201  tcacggtcct caccccatagg agccaagccc agctcatgat gggatcacag ggcagacagc aattcttttt accccggga ctggggccct gggggttgag
6301  gagttggctc tgcagggtct ctaggagagg tggccagatc gcctctgagg ttagagaagg ggaccccttt tacttttcct gaatcagcaa tccgagcctc
6401  cactgaggag ccctctgctg ctcagAACCC CAAAAGGGAG ATTCAAAAGA TCCTGGAGTT TGTGGGGCAC TCCCTGCCAG AGGAGACCGT GGACTTCATG   EXON VII
                                  N  P   K  R  E   I  Q  K  I   L  E  F   V  G  H   S  L  P  E   E  T  V   D  F  M
6501  GTTCAGCACA CGTCGTTCAA GGAGATGAAG AAGAACCCTA TGACCAACTA CACCACCGTC CCCCAGGAGT TCATGGACCA CAGCATCTCC CCCTTCATGA
      V  Q  H  T   S  F  K   E  M  K   K  N  P  M   T  N  Y   T  T  V   P  Q  E  F   M  D  H   S  I  S   P  F  M  R
6601  GGAAAGgtgg gtgctggcca gtacgggggt ttggggcggg tgggagcagc agctgcagcc tccccatagg cactcggggc ctcccctggg atgagactcc
        K  G
6701  agccttgctc cctgccttcc cccccagGC ATGGCTGGGG ACTGGAAGAC CACCTTCACC GTGGCGCAGA ATGAGCGCTT CGATGCGGAC TATGCGGAGA   EXON VIII
                                       M  A  G  D   W  K  T   T  F  T   V  A  Q  N   E  R  F   D  A  D   Y  A  E  K
6801  AGATGGCAGG CTGCAGCCTC AGCTTCCGCT CTGAGCTGTG AGAGGGGCTC CTGGGGTCAC TGCAGAGGGA GTGTGCGAAT CAAACCTGAC CAAGCGGCTC
        M  A  G   C  S  L   S  F  R  S   E  L
6901  AAGAATAAAA TATGAATTGA GGGCCTGGGA CGGTAGGTCA TGTCTGTAAT CCCAGCAATT TGGAGGCTGA GGTGGGAGGA TCATTTGAGC CCAGGAGTTC
7001  GAGACCAACC TGGGCAACAT AGTGAGATTC TGTTAAAAAA ATAAAATAAA ATAAAACCAA TTTTTaaaa gagaataaaa tatgattgtg ggccaggcat
7101  agtggctcat g 7152
```

FIG. 5C

```
-3729  ctctccctcc ttgtctctta cctgcctgct gcctgggaca ggatgaagcg gggcccttgt gttgccccaa ccctggctgt tggctaagag cccacgtgat
-3629  ctgcctgtga gaggagttcc ttccggaaga accagggcag cttctgcccc tagagggcca atgccctagc tgagtgcagt cccccggccc cagcctggtc
-3529  cagctttggg aagagggtgc ccagttgtgc aatccaggcc ggggcagccg tgtcctgatc ttggtattca gggctgagcc tggagggggc ttgtgatgcc
-3429  tgactctgtc tctctctctg gccccatgcc ttggtagctg tgaggcgtca ctgctttggg tgacctgatc tggctgtgat ggatgagcac gggggaaata
-3329  gtggaagact cggaattaga agacgtgagt gggctttggc cccagcctcc ctacccccact ccctgtcctg ggctgcctgt gaccaacctt gtttctgcag
-3229  gcacactgga tagccctgct ggagctcagt gtccctaatc ccctccagat actggtggcc taggggaggt catcaaagac cagtgggaca tcgacctcag
-3129  cctgtttcca cgtttcttgt tgtttttttt tttttgtgga gacagagttt cactcttgtt gcccaggctg gagtgcaatg gcgtgatctt ggctcaccgc
-3029  aacctctgcc tcccgggttc aagcgattct cctgcctcag cctcccaagt agctgggatt acaggcgtgt gccaccaggc ttgactaatt ttctattttt
-2929  agtagagaca aggtttctcc atgttggtca ggctggtctc aaactcccga cttcaggtga tctgcctgcc tcggcctccc aaagtgctgg gattacagga
-2829  gtgagccacc gtgccaggcc ttctccaggc tcttggcacc ttagccagaa acaatttaag gacaagtgca aaagtcatga acgtaggcag atttcctgca
-2729  gagtaaaggg actcactgaa gaagaggaac gtgggggtcc tcaagagagt gtctcatgcc ctacaaggtg tggggctgac ctttatgggc ttcttcaact
-2629  aaagaggggt atattcatga agagtccagg aaaaggtaaa gatttctcaa gaccgtggtg ccacaattta cacccaaata caggtgttcc tggagccgtc
-2529  ttggcactgg tgggtgtacg gtttcatatg ttactgattg tacagtgaga tcctaggtga aacctacatc aaatacagcg ccatgttgct tctggttggt
-2429  cgcagccagc ttggtcctca tcctattttt cagggactta ttggcccttg gcacatgcag ctatttcaag tttccttctt ctggtcatgt gaaactgctg
-2329  cctgggattt tctgttgtct tgctagcact ctattaatct cacattctcg cctcttttct gtgccacccc ctgctggtcc ggctggtttt cactagagtg
-2229  caatacaaag tctcagtcaa gagggcctcc tgaaggttgc tgagggcagg ggtggagcta gtagccggag gacctgccag tcatggggat tcctcagggg
-2129  cacagaggag ggaggagggg cctgtggccc tagcagggga gcagcctctc ctctgcctgg aaatcccatg cctcagtttt ccccgcttgc ctctgagctc
-2029  acgcaaccct gggaaggctt gggagactca cctttactca gatggttgtt tacctgtctc gtgcccaggt tgaccctgga ctttaaatag tgaggacaaa
-1929  gaacgaggag ggtgggggga tgcactcctt ccacgggggc ctgtggcttc caagcctcaa cctcctctgg tctctgtctg tggagcctcc ttcaaaccca
-1829  tggaaagaaa agtacctgcc aggggctgtg gttcttctag gatcttctat cgatgttctg tgaggtcccc agggagccat gaagctgggg ctggctccca
-1729  gggcaatggg actgcagtgt ccttgttctt tcttggttct atggatccat gctctgctcc acccctgccc cttcactctg cccacacgca tcactccaga
-1629  ctggccttgt ggtcagagcc tggagtgcat gggctgctgg aggcctgtgg gttgcactgg gccaggaccc ctggcacctt caagactggc ctggagccag
-1529  caggtaggtg acctttccag ggcctgccta tcccagcttt ctcctccaat ccctcccctc tcttgcctgg gtcaattaga gaaagcttgt cttttggagt
-1429  tcaggggcag gtcaggagcc cagtgacagc tcaaaaaaaa aaccccaaaa aaaaaacccc accattgggc cctttcccct ttcattcttc tgttttctac
-1329  acaccaaacc cagtcgtggc tttggagatc actttaagct tgtctccagc tggcaaacta aggagggtaa tagagaagct cccccacccc caaccctacc
```

FIG. 6A

```
-1229  cottocttcc ggaagcaaat ctaagtccag ccccggctcc agatccctcc cacactgacc taagaaaccc tcagcacaga caacacccct GCATTCCCCA  EXON IB
-1129  CACAACACCC ACACTCAGCC ACTGCGGGCG AGGAGGGCAC GAGGCCAGGT TCCCAAGAGC TCAGgtgagt gacacaccgg aatggcccag gacgccctca
-1029  ccctgctcag cttgtggctc caacattcca gaagccgagg cctctgttat ctctgccctc tccccatgga tatcccattt cagacaaccc cggccggcct
 -929  gaatcccoct cccttccttt ttttttttcc ggggaggcca ggtcttgctg tcaccgaggc tggagtgctg tgggatcctg gccactgcag ccttgaattc
 -829  ctgggctcaa gtgattctcc tgcctcagta gctaggacta cagaccctca ccatcctgcc tggatagttt taaaaaatat ttttaaaaga tttttagaga
 -729  tggggtcttc caatgctgcc cagattggtc tccaaattct ggcctcagcc tccctagggt ctgggattac aggtgggagc caccctgccc aggatcctcc
 -629  ttttgctgag tcatcacagt tttgctcatt ccccatcag gctctggccc ccaataccag ctcagttgct caatgggctG TTTGTCCTGG AACCCAGATG  EXON IA

 -529  GACTGTGGCC GGGCAAGTGG ATCACAGGCC TGGCCAGCCT AGGAGTTGCC ACATGTGAGG GGCCGAGGGG CTCAAGGAGG GGAACATCGG GGAGAGGAGC
 -429  CTACTGGGTG GAGGCTGGGG GTCCCAGCAG GAAATGGTGA GACAAAGGGC GCTGGCTGGC AGGAAGACAG CACAGGAAGG TCCTAGAGGT TCCTCAGTGC
 -329  AGCTGGACTC TCCTGGAGAC CTTCACACAC CCTGACATCT GGGCCCCGTT CCACGAGGGT GCTTTCACTG GTCTGCACCA TGGCCCAGGC CCTGGGATTT
 -229  TGAACAGCTC CGCAGGTGAA TGAAAGgtga ggccaggctg gggaaccacc acattagaac ccgacctggt tttcagcccc agccccgcca ctgactggcc
 -129  ttgtgagtgc gggcaagtca ctcaacctcc ctaggcctca gtgacttccc tgaaagcaag aattccactt tcttgctgtt gtgatggtgg taagggaacg
  -29  ggcctggctc tggcccctga cgcagGAACA TGGAGCTGAT CCAGGACATC TCTCGCCCGC CACTGGAGTA CGTGAAGGGG GTCCCGCTCA TCAAGTACTT  EXON II
                                    M   E  L  I   Q  D  I   S  R  P  P   L  E  Y   V  K  G   V  P  L  I   K  Y  F
   72  TGCAGAGGCA CTGGGGCCCC TGCAGAGCTT CCAGGCCCGG CCTGATGACC TGCTCATCAG CACCTACCCC AAGTCCGGgta ggtgaggagg gccaccccacc
        A  E  A   L  G  P  L   Q  S  F   Q  A  R   P  D  D  L   L  I  S   T  Y  P   K  S  G
  172  ctctcccagg tggcagtccc caccttggcc agcgaggtca tgctcacctc agcctgctca cctcccatct ccctccctct ccagGCACCA CCTGGGTGAG  EXON III
                                                                                               T  T   W  V  S
  272  CCAGATTCTG GACATGATCT ACCAGGGCGG TGACCTGGAA AAGTGTCACC GAGCTCCCAT CTTCATGCGG GTGCCCTTCC TTGAGTTCAA AGTCCCAGGG
        Q  I  L   D  M  I  Y   Q  G  G   D  L  E   K  C  H  R   A  P  I   F  M  R   V  P  F  L   E  F  K   V  P  G
  372  ATTCCCTCAG gtgtgtgtgt cctgggtgca aggggagtgg aggaagacag ggctggggct tcagctcacc agaccttccc tgacccactg ctcagGGATG  EXON IV
        I  P  S  G                                                                                             M
  472  GAGACTCTGA AAAACACACC AGCCCCACGA CTCCTGAAGA CACACCTGCC CCTGGCTCTG CTCCCCCAGA CTCTGTTGGA TCAGAAGGTC AAGgtgagac
        E  T  L  K   N  T  P   A  P  R   L  L  K  T   H  L  P   L  A  L   L  P  Q  T   L  L  D   Q  K  V   K
```

FIG. 6B

```
 572  tgggcacagt ggttcacacc cgcaatctca gtactttggg aggctgaggt gggaagatcc cttgaagcca gaagttccag ataagtctct tccaaaaaaa
 672  aaacttagct gtgcatagtg gtgtgtgcct gtaataccag ttactcagga ggttgaggtg ggaggatcat ctgagcctag gagtttaagg ttacagcgag
 772  ctatgatcac accagtgcac tccaggctgg gtgacagaga aacactgtct caaaaaacga tgaatagaaa gagtgtccca ccagtgcggt ggctcacacc
 872  tgtaattcca gcacttgaag aggctgaggc aggtggatca cctgagacta ggagtttgag atcagcctgg ccaacatggc aaaacccccat ctctactaaa
 972  aatacaaaaa aattagccgg gcatggtggc aggcatctgt aatcccagct acttgggagg ctgaagcagg agaattgctt gaagctggga ggcagaggtt
1072  gtagtcagcc gagacctcac cattgcaccg cagcctggga aacaagagca aaactctgtc tcaaaaaaaa aagaaaaaaa taaaaaagcg gcaggtggca
1172  gggggctggg cctgttgtgg ctcacgcctg taataccagc actttcggag gtcgaggtgg gcagatcacc caaggttagg agtttgagat cagtctggcc
1272  aacatggaga aaccccgtct ctactaaaaa tacaaaaatt agccaggcgt tggggcaggc gccagtaatc ccagctactc gggaggctga ggaaggagaa
1372  tagcttgcac ctgggaggcg gtggttgcag tgagccgaga ttgtgccact gtactccagc ctgggagaca caacgagaca ttgtttcaaa caaaacaaat
1472  aaatatttta aaaggtttgc cacctgggtg gctcaccgct gtaatgccag catttgggga ggccaagatg ggtggaccgc ttgagctcag gagttccaga
1572  ccagcccagg aaacatgggg agactccatc tctataaaag atgcaaataa tcagcagggc atggtggcat agcgctatag tcccagctac tcaaaagtct
1672  aaggttggag gattgcttga gcctgggagg tcaacgttgc agtgagctat tctcactcca gtgcactcca acctgggcaa caggaaaaaa gaaagcccaa
1772  ggtctttttt ctcttttctc ttttttttga gacctagagt cccccccccc aaaaaaaaaa aaaccacaac aaaaagaaaa aagcaaaggt ccaggtgtgg
1872  ggcatgtgaa tccagggaag gaggccccgg ctcagcccag ctttggtcct gttcttctgg gagagtcgcc tcacttcctc cagacttgtc tcatcttcca
1972  cgggggggac tgtctgcctt ttgctctgat gaccaaaaac atgagactct tccgggtaga cctaagaaag gtagagggtg ggtcctcaca gacccacaaa
2072  atttggtggt ggtgggaaca tgcctggtgg agcatgcctt gctccagatc ggggtgtgac gcattgatgc agattatatt actatagaat atgatggtct
2172  cagggaccag gcaggacttt ggcttttgag cagggttcag atcctgactt ggccctacct gtgccgtgag atctcaaaca agtcagcctc taagcctcag
2272  cttcctcctt tgccaaacca agagatgagc tggcctgggg caggctgtgt ggtgatggtg ctggggttga gtcttctgcc cctgcagGTG GTCTATGTTG  EXON V
                                                                                                     V   V Y V A
2372  CCCGCAACGC AAAGGATGTG GCGGTTTCCT ACTACCACTT CTACCACATG GCCAAAGTGT ACCCTCACCC TGGGACCTGG GAAAGCTTCC TGGAGAAGTT
       R  N  A   K  D V    A  V S  Y   Y  H  F   Y H  M   A  K  V  Y   P  H  P   G  T  W    E  S  F  L   E  K  F
2472  CATGGCTGGA GAAGgtgggc ttgatgggag gaaggaaggt gtggagctaa ggggtggtgg ctacaacgca cagcaaccct gtgtcggcac cccctgcccg
       M  A  G   E V
2572  cttctccagT GTCCTATGGG TCCTGGTACC AGCACGTGCA AGAGTGGTGG GAGCTGAGCC GCACCCACCC TGTTCTCTAC CTCTTCTATG AAGACATGAA  EXON VI
                  S  Y  G   S  W  Y  Q   H  V  Q   E  W  W   E  L  S  R   T  H  P   V  L  Y   L  F  Y  E   D  M  K
2672  GGAGgtgaga ccgcctttga tgcttccctc cacgtgacac ctgggggcag gcacttcaca gggacctgcc aaggccaccc agccaccctc cctggcggc
      E
```

FIG. 6C

```
   1 acctctgcct cctggttcca agcaatcctc cttcctcacc ctccagagta gctgggatta
  61 cacgcgcctg ccaccgcgcc tggcctaatt tttgtatttt tagtagagat gggggtttcc
 121 aaccatgttg gccaggctgg tctccaaact cctgacctca ggtgatcctg cccacctaag
 181 cctcccaaaa tgctggtatt acaggcatga gccaccgtgc ccggcctaaa taattaataa
 241 aataatggac gatgggtgcc ttctactgag ctcccggtaa ttgtgagtga gtagaggact
 301 tgccctgggg acattcagtg acctgctggg tgttgctgag ctgtgaggaa gttcaggtct
 361 ggctgcagtg gtgaggctgt gactcaatca atcactgctg atgctcccag gacctgcacc
 421 agcttagtcc taggggcaag gattttaact gtccacctca gtttcttcat ttgtaagatg
 481 caaataacag tcacccctgc ctcatgggat ggagctgtgt aatgcccgca acagtgcctg
 541 ctgcatagag gggttgctgc cagctgcctc tccctccttg tctcttacct gcctgctgcc
 601 tgggtcagga tgaaggaggg ccgttgtgtt gccccccacc tggctgcctg ctaagggccc
 661 atgtgatctg cctggcagag gagtttcttc aggaagaacc agggcagctt ctgcccctag
 721 agggccaatg cccttggtga gtgcagtccc ctggccccag cctggtccac ctctgggaag
 781 agggtgccca gttgtgcaat ccaggcccag gcagctgagc cctcatctca gcatgcaggg
 841 cggatactgg agggggcttg tggcatctga ctctgtatct cctacctgcc cctctccttg
 901 gtagctgtga gaagtcactg ctttggggag acctgatctg gctgtgccag atggacactg
 961 agaaagaagt agaagactca gaattagaag aggtgagtgg gctttggtgg cgggctccct
1021 accccactcc ctgccctggg ctgcctgtga ccacactgct tgcctctgca ggcacactgg
1081 acagacctgc tggagacctg atcctcagtg tccttacccc ctcctacctc ttttctgtgc
1141 cacctgctgt gggtccagca ggtttttact tgagtacaat aaaaagtctg agtcaagggt
1201 gccttatggt ggatgctgag gggaggggcg gagctagtag cccaaggtcc tgccagtcac
1261 ggggcttcct caggggcaca gaggaggcag gaggggcccc tggccctagc acgtgaacag
1321 cttctactct gcctggaaac cccatgcctc agctttcccc tacttgcctc tgagctcatg
1381 caattcttgg aagcctggga gacttacctt gaaattgaat gcaaatagga caaagaccaa
1441 ggaggatggg gggatgccct ccttccacgg ggccctgtgg cttccaagtc ttaatctcct
1501 ctagtctctt gtctacggag cctccttcaa acccagggaa agaaaagcac ctgccagggt
1561 tgtttttctt ctaggatctt ctattgatgc tctgtgaggt cccccaggag ccatgaagct
1621 agggctggct cctagggcaa tgggactaca gtgtccttgt cctttcttat tctttctgtt
1681 ctttctttct ttctttttt ttttttttt ttttttgag acagagtctc actctgttgc
1741 ccaggctgga gtgcagtggt gtgatcttgg ctcactgaaa cctccgcctc ctgggttcaa
1801 gtgattctct tgcctcagcc tcctgagtag ctaggattac aggtgcccgc catcatgccc
1861 agctaatttt tgtatttta gtagagacag ggtttcacca tgttggccag cttggtctcg
1921 aactcctgac ctcaggtgat cctgctgcat cgacctccca aagtactggg attacaggcg
1981 tgagccacca cgctcagcct ctttcttgtt ctatatgtcc atgctctgct ccacttctgc
2041 cccttcactc tgccccacac atcactccag actggccttg tggtcagagc ctggaatgcc
2101 tgggctgctg ggggcctgtg gactgcactg ggccagaacc cctgccgcct tcaagactgg
2161 cctgtagcca gcaggtaggt gacttttccc aggccggcct atcccacctt tcccctccac
2221 tcactcacct cccttgcctg ggtcaattag agaaagcttg tcggccaggc atggtggctc
2281 atgcctgtaa tctcagcact ttgggaggcc gaggcgggcg gatcatctga gctcaggagt
2341 ttgagaccag cctggccaac atggcaaaac cccgtctcta ctaaaaatac aaaaattaac
2401 cggatgtggt ggtgtgcacc tgtaatccca gctactcggg aggctgaggc agaagaatcg
2461 cttgaaccca ggaggggggag gttacagtga gcggagatcg tgctactgca ttgcagcctg
2521 ggcgagagag cgagtctcca tctcacataa aaaaaagaaa aagaaagaaa gcaagcttgt
2581 ctgttggcct gccctgcagg gtggagttca gagggaaggt caggagccta gtgacagctc
2641 aaaaaaaaaa aaacccaaat accaatgttg gccccttttg cctttcattc atgtgttttc
2701 tatacactaa actcacatat tgggtttgca gatcactcca agcttggctg gagctgtggt
2761 ggtaaggagg gtaatagaga agcttcccca ccctcaaccc caccccttcc ttcctggagt
2821 tcccagccct gactttagat ccctccaca ctggaccttc aaaaccctca gggcagagag
2881 cagccctaca ctccctacac cacacccata ctcagcccct gcaggcaagg agagaacagg
2941 tcaggttccc gagagctcag gtgagtgaca cgttggaatg gcccagggca ccttcaccct
3001 gctcagcttg tggctccaac attctagaag ccgaggcctc tgccatccct gccctttccc
3061 atggatattc catttcaatt agacaaccca gcctggccgg aatccccctg cgttccttct
3121 tttcctttgt gtatttttga gacagggtgt tgctccgtca cccaggctgg agtgtagtgg
3181 gatcctggcc cactgcagcc tcaaattcct aggctgaggc aatcctgccg cctcagcctc
```

FIG. 7A

```
3241 ctgagtagct ggggttacaa gagcaagcca ccacacccag ctaattttga aaaatatttt
3301 ttgtagagga gaggtcttgc tttgttgtcc aggttggtct caaactccag ggctcaaggg
3361 atcctttccc gttggcctcc caaggctctg ggattacagg cgggagtcac cctgcctggg
3421 cccctccttt tgatgagtca tcagttttca ttcccgcacg aggctctagc ccctggtacc
3481 agcttagttg ctcaatgggc tgtgtttgtt ctggagccca gatggactgt ggccaggcaa
3541 gtggatcaca gacctggccg gcctgggagg tttccacatg tgaggggcat gaggggggct
3601 caaggagggg agcatcgggg agaggagcgc actgggtgga ggctgggggt cccagcagga
3661 aatggtgaga caaagggcgc tggctggcag ggagacagca caggcaggcc ctagagcttc
3721 ctcagcacag ctggactctc ctggagacct tcacacaccc tgatatctgg gccccgcgct
3781 acgagggtgc tttcactggt ctgcactatg ccccaggccc tgggattttg aacagctctg
3841 caggtgactg aaaggtgcgg ccaggctggg gaacgacctg gttttcagcc cagccccgcc
3901 actgactgac tttgtgagtg cgggcaagtc actcagcctc cctaggcctc agtgacttcc
3961 ctgaaagcaa aaactctgca aaggggcagc tgggtgctgg ctcacacctg taatcccagc
4021 actttgggag gctgaggtag acaaatcact tgaggccagg agttctagac cagcctggcc
4081 aacatggtga aaccccatct ctactaaaga aaaaaaaaaa ttagctgagc atggttgtac
4141 atgcttgtaa tcccagctac ttgggatgcc gaggcgggag gattgcttga acccaagagg
4201 tggagtttgc agtgagctga gattgtgcca cactgcactc cagcttgggt gagagtgaga
4261 ctccatctca aaaaaaaaaa aaaaaagaga gaatcccact ttcttgctgt tgtgatggtg
4321 gtaagggaac gggcctggct ctggcccctg atgcagaaac atggagctga tccaggacac
4381 ctcccgcccg ccactggagt acgtgaaggg ggtcccgctc atcaagtact ttgcagaggc
4441 actggggccc ctgcagagct tccaagcccg acctgatgac ctgctcatca acacctaccc
4501 caagtctggt aagtgaggag ggccacccac cctctcccag gcggcagtcc ccaccttggt
4561 cagcaaggtc gtgccctcag cctgctcacc tcctatctcc ctccctctcc acgcaccacc
4621 tgggtgagcc agatactgga catgatctac cagggcggcg acctagagaa gtgtaacctg
4681 gctcccatct acgtacgggt gcccttcctt gaggtcaatg atccagggga accctcaggt
4741 gcatggctgg gtcctggggg taagggaagt ggaggaagac agggctgggg cttcagctca
4801 ccagaccttc cctgacccac tactcagggc tggagactct gaaagacaca ccgccccccac
4861 ggctcatcaa gtcacacctg cccctggctc tgctccctca gactctgttg gatcagaagg
4921 tcaaggtgag gccgggctca atggttcaca cctgtcatcc cagtttgaga ctgaggaggg
4981 aggatccctt gaaggcgaga gatggagacc agcctgggca acattgctgt agagatgaca
5041 tcccatctct acaaaaataa aattaacaac ctggtatggt ggcatagact gttcccagtt
5101 acttaggagg ctcagcgggg aggactgttt atgcaaatag gaagctgcaa tgagccctga
5161 tgatcctgct gctgcactcc agcctgggca acacagcaaa accatctcta cgaaaaaaaa
5221 agttcccact gactggcaag gaaagccagg aaggggggct caggtgccct ctcagccatg
5281 tacctgttct tctggaaggg cctcctcgct tctgccaggc tcatcacatc ttttttttt
5341 ttgagacaga gtcttgctct gtcaccctgg ctggagtgca gtggcatgat ctcagctcac
5401 tgcaacctcc gcctcccag ttcaagtgat tctcctgcct cagcctcctg agtagctggg
5461 attacaggcg tgtgctacca cacccggcta atttttgtat tcttttagt agagacgggg
5521 tttcaccatg ttggtcaagt ggatctcaaa ctcttgacct tgtgatcctc ctgcctcgac
5581 ctcacaaagt gctggaatta caggcgtgag ccaccgcgcc tggcctttt ttttttgag
5641 acagtttcac tcttgttgcc gaggctagag cgcaatcgtg tgatctcggt tcactgcaac
5701 caccgcctcc tgggttcaag caattctcct gcttcagcct cccaaggagc tgggattaca
5761 ggtacctgcc accacgcccg gctaattttg tatttttagt agagatgggg tttcaccatg
5821 ttggtcaggc tggtcttgaa ctcctgacct caggtgatct ggcaccttgg cctcccaaag
5881 tgccgggatt agaggcatga gccaccacgc ccagccttca tcacatcttg agagaggaca
5941 ctgtctgcct cttgctctga tgagggtctg atgcaaagga tagtgagtct ctacagtgca
6001 cacttaagaa aggcagcatg tgggtgctca caggtcaggc ggaggagggg gagctggtgg
6061 ggaccaggca tgccttgctc cagatcagga tatgatggca ttggtgcaga ttatattagt
6121 atagaatatg gtctcaggaa ccaggcagga ctttggcttc cgagcagggt tcagatccca
6181 gcttggccct acctgtgcag tgagatctca agcaagtcag cctctaagcc tcaggttcct
6241 cctttgccag ttcaacagat gagctggcct ggggtgggct gtgtggtgat ggtgctgggg
6301 ctgggtcctc tgccccctgca ggtggtctat gttgcccgaa acccaaagga cgtggcggtc
6361 tcctactacc atttccaccg tatggaaaag gcgcaccctg agcctgggac ctgggacagc
6421 ttcctggaaa agttcatggc tggagaaggt gggcttgact ggaggaagga gggtgtgaag
6481 ccgaggggtg gtggctataa cgtacagcaa ccctgtgtcg gtgccccctg cccgcttctc
6541 tagtgtccta cgggtcctgg taccagcacg tgcaggagtg gtgggagctg agccgcaccc
6601 accctgttct ctacctcttc tatgaagaca tgaaggaggt gagaccgact gtgatgcttc
```

FIG. 7B

```
6661 cccccatgtg acacctgggg gcaggcacct cacagggacc caccaaggcc acccagcccc
6721 gtccctgggc ggctcccaca gcaagcccgg attccccatc ctacctccct ggcccaggcc
6781 cccccactgc agccccacct ggcagcaggc tcggcacagc tttcatcttc tgcacctgag
6841 tcagctgcat gggtggccac ggatcagata cttagtccta ttgcttatcc tcaccaaagg
6901 gtgtgccacc cagggccaca gtcatggaag aagaccatcc cggtcctcac ccataggcgc
6961 caagccctgt tcatgatggg atcacagggc agagatcaat tcattttact ccaggacta
7021 gggccccagg ggttgaggct ctttggggtt tctaggggaa gtggccagat cccctctgag
7081 gttagagagg gggacccgtt ttgttttgct ccactgagga gccctctgct gctcaggaacc
7141 ccaaaaggga gattcaaaag atcctggagt ttgtggggcg ctccctgcca gaggagacca
7201 tggacttcat ggttcagcac acgtcgttca aggagatgaa gaagaaccct atgaccaact
7261 acaccaccgt cccccaggag ctcatggacc acagcatctc ccccttcatg aggaaaggtg
7321 ggtgctggcc agcacggggg tttggggcgg gtgggagcag cagctgcagc ctccccatag
7381 gcacttgggg cctcccctgg gatgagactc cagctttgct ccctgccttc ctccccccag
7441 catggctggg gactggaaga ccaccttcac cgtggcgcag aatgagcgct tcgatgcgga
7501 ctatgcggag aagatggcag gctgcagcct cagcttccgc tctgagctgt gagaggggct
7561 cctggagtca ctgcagaggg agtgtgcgaa tctaccctga ccaatgggct caagaataaa
7621 gtatgatttt tgagtcaggc acagtggctc atgtctgcaa tcccagcgat ttgggaggtt
7681 gagctggtag gatcacaata ggccacgaat ttgagaccag cctggtaaaa tagtgagacc
7741 tcatctctac aaagatgtaa aaaaattagc cacatgtgct ggcacttacc tgtagtccca
7801 gctacttggg aagcagaggc tggaggatca tttcagccca ggaggttgtg gatacagtga
7861 gttatgacat gcccattcac tacagcctgg atgacaagca agaccctccc tccaaagaaa
7921 ataaagctca attaaaataa aatatgattt gtgttcatgt agagcctgta ttggaaagga
7981 agagaaactc tgagctgaaa gagtgaatgc ccggtggggc cacatatggt cacctctccc
8041 ccagccttca gctccccagg tcaccatatc tggggagggg agaagggttt ggagaagtaa
8101 aacccaggag atgtgtggag gggggatgtc tgtttaatcc cagcacatcc tctgctgtcc
8161 tgccccaaga tggtggagga cgtcgagtcc gccgggcagc gtcacttttt cttgggctcc
8221 ttagaagcta ccaggtacct ctgggccaca ctgagatgag gggagtagcc gcctgcatag
8281 gaggtgtctt caaacaggat agtatagtcc ctcctggggg ttgtgggggt aggtggccaa
8341 ggaagggtag aggagcaagc ccccggggct ggttgtcaac tcactttgtt ggctggaatt
8401 ggttgtaact tgaccacctc gggcaggatc ccactgctca tccccaa
```

FIG. 7C